EDGEWOOD

CHEMICAL BIOLOGICAL CENTER

U.S. ARMY RESEARCH, DEVELOPMENT AND ENGINEERING COMMAND

ECBC-TR-534

NEW MODEL FOR POPULATION-SUBPOPULATION DIFFERENCES

Ronald B. Crosier

RESEARCH AND TECHNOLOGY DIRECTORATE

February 2007

ABERDEEN PROVING GROUND, MD 21010-5424

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
XX-02-2007	Final	Nov 2005 - Nov 2006

4. TITLE AND SUBTITLE
New Model for Population-Subpopulation Differences

5a. CONTRACT NUMBER
DE-AI01-00NN20105

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Crosier, Ronald B.

5d. PROJECT NUMBER
01-01NN20105.001

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DIR, ECBC, ATTN: AMSRD-ECB-RT-DD, APG, MD 21010-5424

8. PERFORMING ORGANIZATION REPORT NUMBER
ECBC-TR-534

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DOE, ATTN: NNSA NA-22 (Dr, Terry Creque),
1000 Independence Avenue, Washington DC 20585

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
DOE TPOC: Sandia Corporation dba Sandia National Laboratories, ATTN: John E. Brockmann, Mail Stop 0836, PO Box 5800, Albuquerque, NM 87185-0836

14. ABSTRACT
Civil defense planning requires estimates of the toxicity of chemical warfare agents to the general public, but the current toxicity estimates are for male soldiers. In ECBC-TR-224 and ECBC-TR-337, individual susceptibilities for both the general population and the military subpopulation were modeled by a lognormal distribution. The assumption of a lognormal distribution of individual susceptibilities for both the general population and a subpopulation cannot be correct. This report presents an alternative model and compares the previous and new models.

15. SUBJECT TERMS

Lognormal distribution	Risk assessment	Normal distribution	Demographics
Conversion factor	Subpopulation	Chemical warfare	Probit slope

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Sandra J. Johnson
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UL	19	(410) 436-2914

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Blank

PREFACE

The work described in this report was authorized under Intra-agency Agreement, Contract No. DE-AI01-00NN20105 and Project No. 01-01NN20105.001 with the U.S. Department of Energy, National Nuclear Security Administration, Chemical and Biological National Security Program. The work described in this report was started November 2005 and completed in November 2006.

The use of either trade or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.

This report has been approved for public release. Registered users should request additional copies from the Defense Technical Information Center; unregistered users should direct such requests to the National Technical Information Service.

Acknowledgments

The author thanks the editor and reviewers of The American Statistician for helpful comments, as well as the anonymous associate editor of The American Statistician for suggesting the new model described in this report.

Blank

CONTENTS

FIGURES

TABLES

NEW MODEL FOR POPULATION-SUBPOPULATION DIFFERENCES

1. INTRODUCTION

Civil defense planning requires estimation of casualties from the use of chemical warfare agents against the civilian population. Computer models of atmospheric transport and dispersion can estimate the exposure of the civilian population to chemical warfare agents from a given scenario. To assess casualties, these models require estimates of the toxicity of chemical warfare agents to the general population. The available estimates of the toxicity of chemical warfare agents (Grotte and Yang 2001) are for military personnel. It is widely believed that the general population is more susceptible to toxicants (harmful substances) than the military subpopulation is and that the general population has more variability in susceptibility to toxicants than the military subpopulation does. In the absence of data relevant to the soldier-to-civilian adjustment, a subjective estimate must be used. A common practice in toxicology is to account for an unknown difference by applying an uncertainty factor; the default uncertainty factor for the difference between a population and a subpopulation is 10—see, for example, Whalan, Foureman, and Vandenberg (2006). Uncertainty factors are typically applied to a low percentile of a distribution to estimate a safe level of exposure. Application of uncertainty factors to the parameters of a distribution is unusual, but given the lack of methods for converting the parameters for military personnel to parameters for the general public, such an application might be made. Concern that estimates based on a factor of 10 might exceed what is mathematically possible led to Crosier and Sommerville (2002) and Crosier (2003). The model used in those reports has been criticized for its distributional assumptions. This report compares the previous work to a more realistic model that was proposed by an associate editor of a journal.

2. BACKGROUND AND NOTATION

For each individual there is a dose that is just sufficient to cause a specified response. These just sufficient doses are called effective doses to distinguish them from the administered doses of a toxicological study, or the actual doses received by individuals. In toxicology, a dose is an amount, such as two pills, a teaspoonful, or five milligrams. A dosage is an amount relative to something else, such as two pills per day, a teaspoonful with each meal, or five milligrams of a substance per kilogram of body mass. For exposure to toxicants in the atmosphere, the dose (amount absorbed) is unknown. The toxicity of inhaled toxicants is characterized by the exposure concentration and the exposure duration, which can be combined into a single number by one of several models. The distinction among dose, dosage, and exposure is not needed for modeling subpopulations; henceforth, the term dose will be used generically for dose, dosage, or exposure. A lognormal distribution of effective doses is typically used both by toxicologists for the analysis of data and by modelers for the prediction of casualties—see, for example, Cornwell and Marx (2006).

Toxicologists characterize a lognormal distribution of effective doses by its median effective dose (ED_{50}) and its probit slope. The probit slope is the reciprocal of the standard deviation of log(effective dose), where log is the common (base 10) logarithm. Therefore, the probit slope has units of standard deviations per one base-10 logarithm unit, or, equivalently, standard deviations per a factor-of-10 change in the dose. For both the population and a subpopulation, these toxicological parameters are related to the mean μ and standard deviation σ of a normal distribution by

$$\mu_p = \log(\text{population } ED_{50}) \quad (1)$$
$$\mu_s = \log(\text{subpopulation } ED_{50} \quad (2)$$
$$\sigma_p = 1 / (\text{population probit slope} \quad (3)$$
$$\sigma_s = 1 / (\text{subpopulation probit slope}) \quad (4)$$

in which the subscripts p and s represent population and subpopulation, respectively.

3. CRITICISM OF THE ASSUMPTIONS

The inconsistency between the assumptions that $\mu_p < \mu_s$, $\sigma_p > \sigma_s$, and lognormal distributions of effective dose for both the population and the subpopulation can be illustrated by a plot of percent of individuals responding to a dose versus the dose on lognormal probability paper. Figure 1 shows the lines for a population with $\mu_p = 0$, $\sigma_p = 0.3$ (solid line) and a subpopulation with $\mu_s = 0.2$, $\sigma_s = 0.1$ (dashed line).

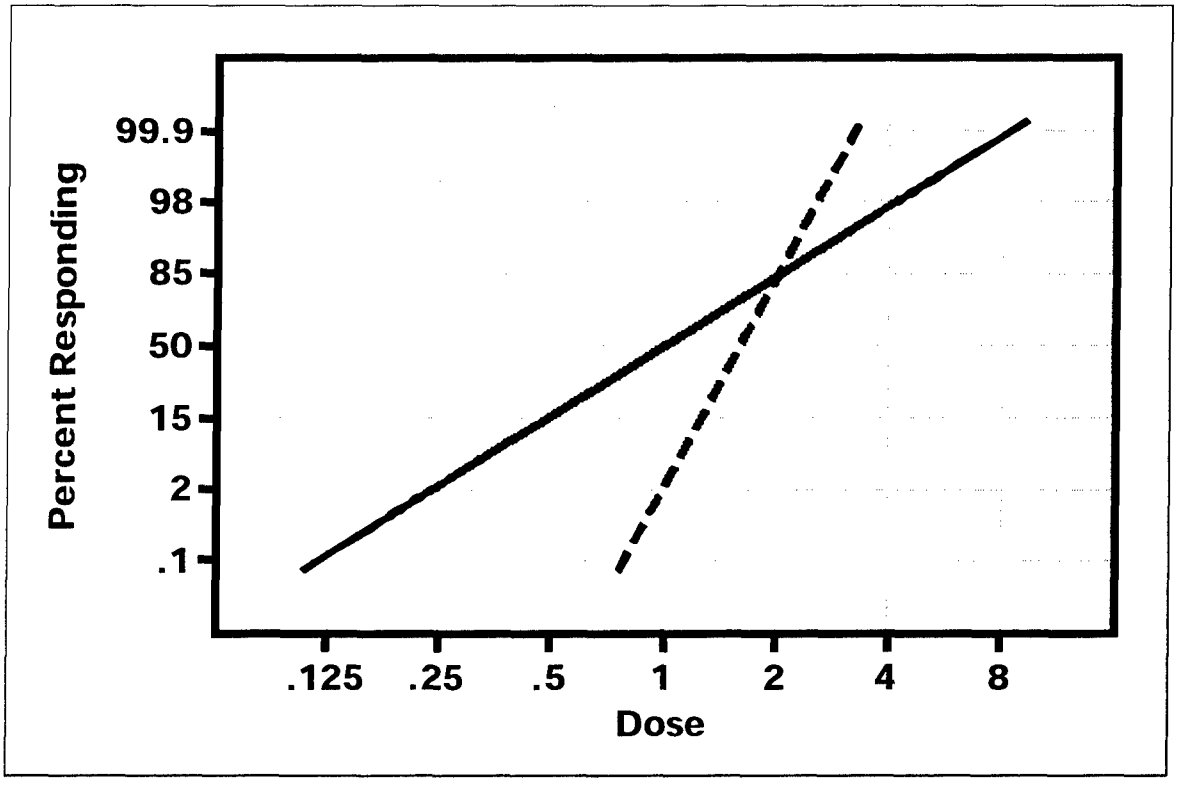

Figure 1. Probability Plot for a Population (solid) and a Subpopulation (dashed)

The lines in Figure 1 cross at dose = 2; at any dose > 2, the fraction of the military subpopulation responding to the dose will be larger than the fraction of general population responding to the dose. This nonsensical result is a direct consequence of the seemingly reasonable assumptions made in the problem formulation.

4. SUBPOPULATION MODEL

This model from Crosier and Sommerville (2002) and Crosier (2003) accepts the assumptions of the problem as given—that is, lognormal distributions of effective doses for both the population and the subpopulation—and checks whether the values for μ_p, σ_p, μ_s, and σ_s are mathematically consistent with those assumptions and with the subpopulation size. Its development is repeated here to establish some concepts and notation and for comparison to an alternative model in the next section.

Figure 2 shows histograms of log(effective dose) for a population and a subpopulation that is 30% of the population. The curves in Figure 2 are not probability densities but frequencies—normal curves fit to histograms—as described, for example, in Dixon and Massey (1969). Letting x = log(effective dose) and h(x) = height of the normal curve fit to a histogram, the equations for the normal curves for the population and subpopulation are:

$$h_p(x) = [\, N_p\, w / \sigma_p\, (2\pi)^{1/2}\,] \exp[-(x - \mu_p)^2 / 2\,\sigma_p^2\,] \quad (5)$$

$$h_s(x) = [\, N_s\, w / \sigma_s\, (2\pi)^{1/2}\,] \exp[-(x - \mu_s)^2 / 2\,\sigma_s^2\,] \quad (6)$$

where N_p is the size of the population, N_s is the size of the subpopulation, and w is the width of the class intervals, or frequency bins, used to construct the histograms.

In Figure 2, the same set of class intervals were used to construct both the population histogram and the subpopulation histogram. Because members of a subpopulation are also members of the population, the subpopulation cannot have more members in a class interval, or bin, of the histograms than the population does. In terms of the normal curves, the height of the subpopulation curve cannot exceed the height of the population curve at any value of x. Consider two cases.

Case 1: Supposing $\mu_s = \mu_p$, we have $h_s(x) \le h_p(x)$ at $x = \mu_s = \mu_p$, immediately leading to $N_s / \sigma_s \le N_p / \sigma_p$, or $(N_s / N_p)\, \sigma_p \le \sigma_s$—that is, the subpopulation standard deviation cannot be less than N_s / N_p times the population standard deviation. It is convenient to define the subpopulation size as a fraction, $\theta = N_s / N_p$, of the population size.

Case 2: Again supposing $\mu_s = \mu_p$, consider how large σ_s may be; σ_s cannot exceed σ_p because the heavier tail of the subpopulation curve would become higher than the tail of the population curve at some large value of x. Thus, the limits on σ_s are established: $\theta\, \sigma_p \le \sigma_s \le \sigma_p$.

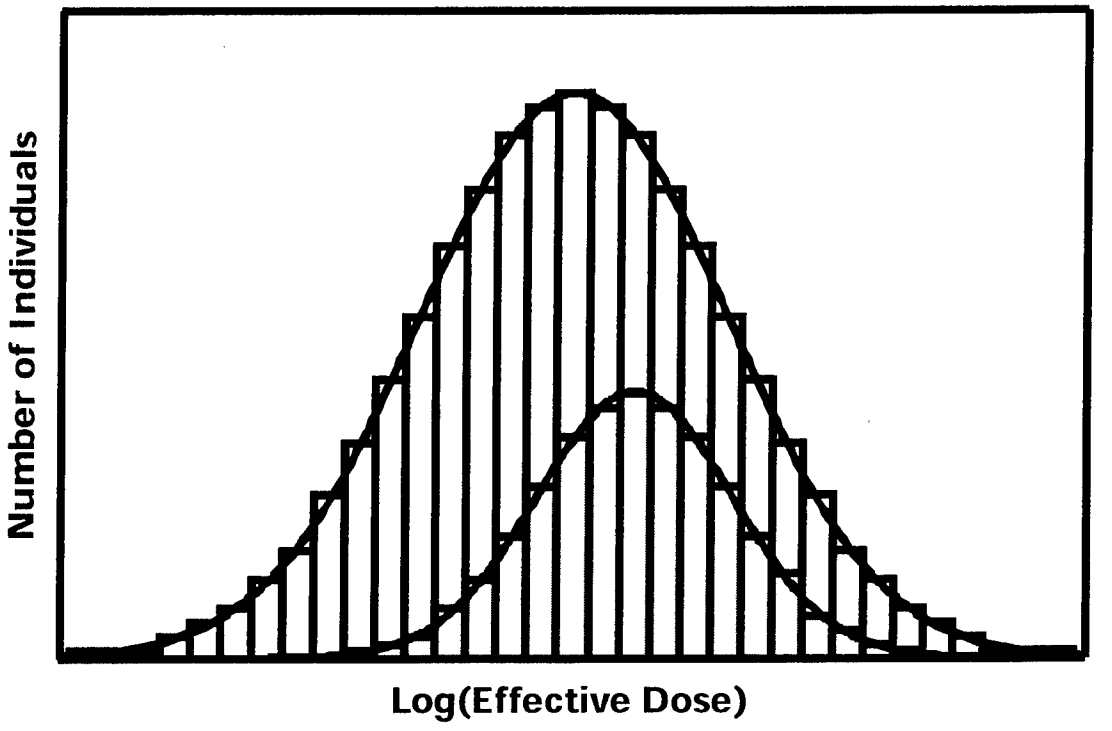

Figure 2. Histogram of Log (Effective Dose); Subpopulation Size is 30%.

For $0 < \theta < 1$ and $\theta\,\sigma_p < \sigma_s < \sigma_p$, μ_s may vary over some range without the height of the subpopulation curve exceeding the height of the population curve at any value of x. If μ_s is equal to a limit of its allowable range, the subpopulation curve will make contact with the population curve at the contact point. Figure 3 uses the normal curves without the histograms to illustrate the contact point for the case with $\mu_s > \mu_p$. The case with $\mu_s < \mu_p$ would yield the mirror image of Figure 3. Denote the x coordinate of the contact point by x_0. At the contact point, the heights of the two curves are the same, $h_s(x_0) = h_p(x_0)$. Also at the contact point, the derivatives of the two curves must be the same, as otherwise the curves would cross. These two conditions (on heights and derivatives) yield two equations that can be solved to obtain an expression that identifies the feasible combinations of the parameters. To simplify the derivation, which is given in Crosier (2003), it is helpful to use the linear transformation $z_p(x) = (x - \mu_p) / \sigma_p$. For the subpopulation, the mean and standard deviation of $z_p(x)$ are

$$\delta = (\mu_s - \mu_p) / \sigma_p \tag{7}$$

and

$$\varepsilon = \sigma_s / \sigma_p \tag{8}$$

respectively, whereas for the population, $z_p(x)$ has mean zero and variance one. The derivation yields

$$\delta = \pm [\, 2\,(\varepsilon^2 - 1)\, \mathrm{LN}(\theta/\varepsilon)\,]^{1/2} \qquad (9)$$

where LN is the natural (base *e*) logarithm. Equation (9) gives the limits of the feasible range for δ as a function of θ and ε. The ranges of θ and ε are $0 < \theta \leq \varepsilon \leq 1$.

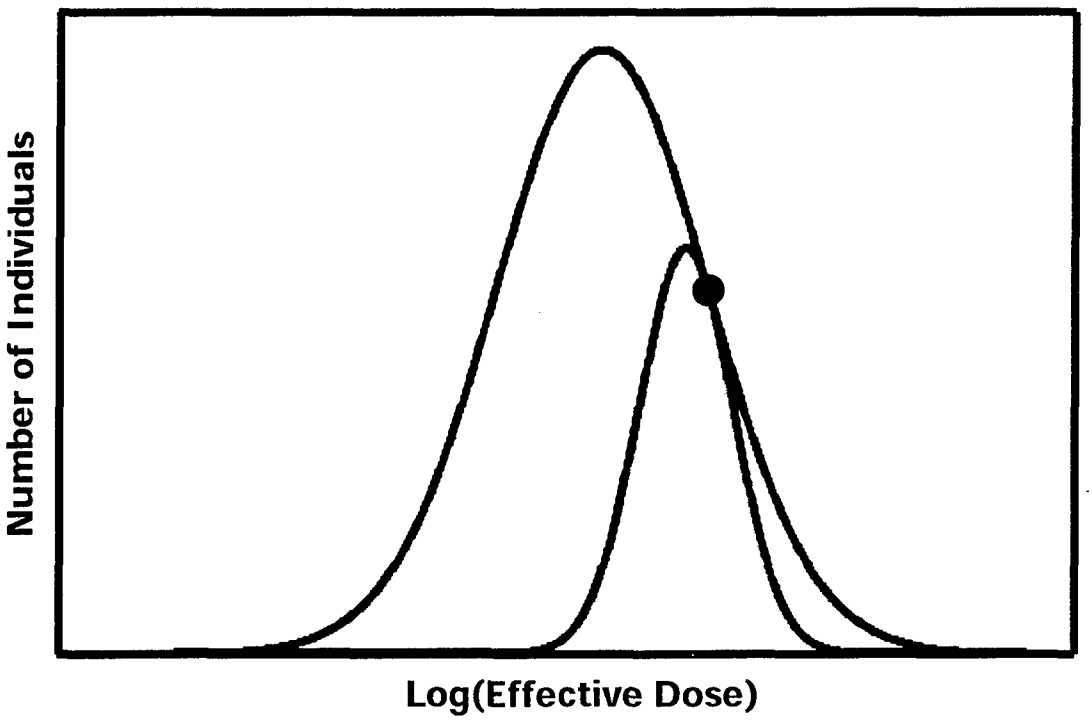

Figure 3. Contact Point of Population and Subpopulation Curves

In toxicological applications, a subpopulation with $\delta > 0$ is called a resistant subpopulation and a subpopulation with $\delta < 0$ is called a sensitive subpopulation. The term feasible region will be limited to either the resistant subpopulation case or the sensitive subpopulation case. It is only necessary to study one case; the results apply to the other case by symmetry. Figure 4 shows the feasible region (shaded) for a resistant subpopulation of size $\theta = .3$. The feasible region appears to be a semi-ellipse, but it is slightly asymmetric in the left-right direction; this asymmetry is more pronounced for smaller values of θ. The line drawn from the origin tangent to the feasible region in Figure 4 touches the feasible region at the point where the ratio δ/ε is maximized. This combination of δ and ε, which is marked by a diamond in Figure 4, yields the minimum value for μ_p because, from (7) and (8),

$$\mu_p = \mu_s - (\delta/\varepsilon)\, \sigma_s \qquad (10)$$

Equation (10) allows calculation of the median effective dose for the general population, which is antilog(μ_p), from the known quantities μ_s, σ_s, and parameters (δ and ε) that describe the relationship between the population and the subpopulation.

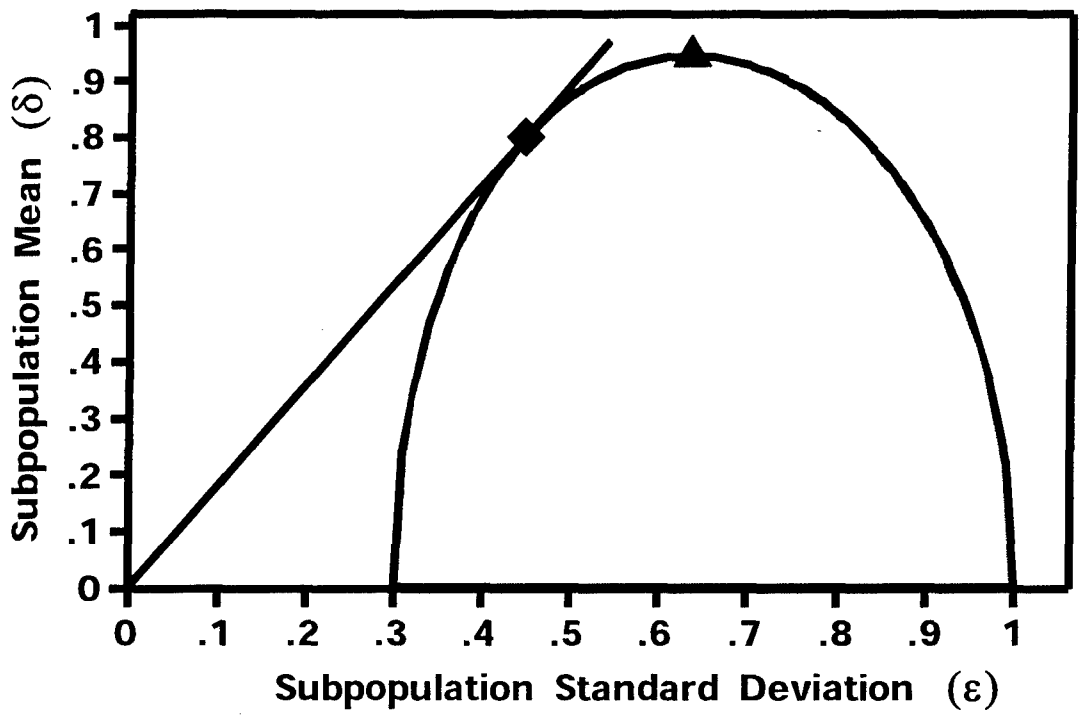

Figure 4. Feasible Region of δ and ε for a Subpopulation of Size $\theta = .3$

The values of ε and δ that maximize the ratio δ/ε may be found analytically. The ratio δ/ε can be at its maximum only if δ is at its maximum for the given value of ε. Therefore, δ in the ratio δ/ε can be replaced by the right-hand side of (9). Monotonic transformations are often used to simplify the process of finding maxima; here, squaring works well. Taking the derivative of $(\delta/\varepsilon)^2$ with respect to ε, setting the derivative to zero, and solving for ε yields

$$\varepsilon_r = \theta \exp[(1 - \varepsilon_r^2)/2] \qquad (11)$$

Setting the derivative to zero fixes the value of ε, so it is denoted ε_r to indicate that it is the value of ε that maximizes the ratio δ/ε. For fixed θ, (11) can be solved by a numerical procedure, then (9) can be used to obtain δ_r, the value of δ that corresponds to ε_r and hence maximizes the ratio δ/ε, from ε_r.

The application is a subpopulation-to-population problem—given θ, μ_s, and σ_s, how large can $|\mu_s - \mu_p|$ be? The solution involves maximizing the ratio δ/ε for fixed θ. A population-to-subpopulation problem—given θ, μ_p, and σ_p, how large can $|\mu_s - \mu_p|$ be?—requires finding the maximum value of δ for fixed θ (the triangle in Figure 4). The maximum value of δ, δ_x, and the value of ε at which it occurs, ε_x, can be found analytically by a procedure similar to the procedure used to find δ_r and ε_r.

5. COVARIATE MODEL

The covariate model, which was proposed by an anonymous associate editor of *The American Statistician*, describes the physical process of generating a subpopulation. Military personnel are selected on the basis of characteristics that may be correlated with effective dose. Let X be the random variable of interest—here, log(effective dose)—and Y be the random variable by which the selection of subpopulation members is made. For convenience, assume that both X and Y are standardized so that the population has mean zero and variance one for both X and Y. If the subpopulation consists of all individuals for which $a \le Y \le b$, then the probability density function (pdf) of X for the subpopulation is

$$f_{Xs}(x) = \frac{\int_a^b f_{X,Y}(x,y)\,dy}{\int_a^b f_Y(y)\,dy} \qquad (12)$$

where $f_{X,Y}(x,y)$ is the joint probability density function of X and Y for the population and $f_Y(y)$ is the marginal probability density function of Y for the population. Note that the integral in the denominator of (12) yields the value of θ. If $a < b$ and $f_{X,Y}(x,y)$ is a bivariate normal distribution with correlation coefficient $\rho > 0$, then $f_{Xs}(x)$ is not a normal distribution. Therefore, the subpopulation model, which assumes that the subpopulation has a normal distribution, cannot be correct for the application. It can, however, be a useful approximation if the distribution of the subpopulation is close to normal. The pdf $f_{Xs}(x)$ can be obtained by numerical integration and compared to a normal distribution with the same mean and variance as $f_{Xs}(x)$ [the subpopulation mean δ and variance ε^2 are obtained numerically from $f_{Xs}(x)$]. I compared $f_{Xs}(x)$ of the covariate model to its normal approximation by using the cumulative distribution function (CDF), $F_{Xs}(x)$. The CDF $F_{Xs}(x)$ gives the fraction of individuals in the subpopulation responding to a dose x. Let *C* be the maximum value of $|F_{Xs}(x) - \Phi((x - \delta)/\varepsilon)|$ for any x, where Φ is the standard normal CDF. Numerical calculations show that, for $\theta > .001$, the criterion *C* does not exceed .01 if $|\rho| \le .69$.

To make a figure showing the feasible region of δ and ε for the covariate model, it is necessary to fix θ and compute δ and ε numerically for various values of ρ, *a*, and *b*. Attention to several special cases may enhance understanding of the covariate model. First, if $\rho = 0$, the selection by Y is irrelevant; the subpopulation has a normal distribution with the same parameter values as the population. Therefore, the case $\rho = 0$ is represented by the single point $\delta = 0$ and $\varepsilon = 1$. Second, if $\rho = 1$, X has the same distribution as Y, which, for the subpopulation, is a truncated normal distribution with *a* truncating a fraction q_1 and *b* truncating a fraction q_2. Johnson and Kotz (1970) give formulas for the mean and variance of truncated normal distributions. Third, if $q_1 = q_2$, then the distribution of X for the subpopulation is symmetrical, $\delta = 0$, and the value of ε depends on the two parameters ρ and $\theta = 1 - q_1 - q_2$. Fourth, if either $q_1 = 0$ or $q_2 = 0$—indicating one-sided or single truncation—the results can again be expressed in terms of the two parameters ρ and θ. If Y represents the health status

of individuals, and the military does not reject anyone for being too healthy, then $q_2 = 0$ is reasonable.

For the general case of single or double truncation, Figure 5 shows the feasible region of ε and δ for $\theta = .3$ from the covariate model (dashed curves and dash-dot line). The apex of the feasible region from the covariate model, which is marked by a square in Figure 5, represents the point $\rho = 1$ and $q_2 = 0$. The parameter combination $\rho = 1$ and $q_2 = 0$ puts the subpopulation into the upper tail of the population and therefore was called the tail model by Crosier and Sommerville (2002). The left side—which has long dashes—represents $\rho = 1$; q_2 varies from zero at the top to $(1 - \theta)/2 = .35$ at the bottom. The right side—which has short dashes—represents $q_2 = 0$; ρ varies from one at the top to zero at the bottom. For single truncation ($q_2 = 0$), the feasible region consists only of this curve of short dashes. The line at the base of the feasible region—which has a dash-dot pattern—represents symmetrical truncation ($q_1 = q_2$); ρ varies from one at the left end to zero at the right end. The upper boundary of the feasible region from the subpopulation model is outlined in gray, and its maximum-ratio and maximum-mean points are again marked with a diamond and triangle, respectively. The circle in Figure 5 marks the intersection of the upper boundary of the feasible region of the subpopulation model with the feasible region of the covariate model for single truncation. This point is the worst-case parameter combination for a subpopulation created by single truncation on a covariate, given the requirement for approximately normal distributions for the population and the subpopulation. The centroid from the subpopulation model (see Crosier 2003) is marked by an asterisk in Figure 5. The feasible region of the subpopulation model indicates approximately the region where the distribution of a subpopulation obtained by truncation on a covariate can be adequately represented by a normal distribution.

Table 1 gives the standardized parameters δ and ε for resistant subpopulations and the goodness-of-fit criterion C for the normal approximation of the subpopulation distribution. The worst-case parameter sets (in terms of maximizing $|\mu_p - \mu_s|$) are given for (1) double truncation, population-to-subpopulation conversion, (2) double truncation, subpopulation-to-population conversion, and (3) single truncation. For single truncation, there is no difference between maximizing δ and maximizing the ratio δ/ε; hence, there is only one worst-case set of parameters for both population-to-subpopulation and subpopulation-to-population conversions. In Figure 5, the three cases of Table 1 are marked by a triangle, a diamond, and a circle, respectively.

The values of θ, δ, and ε in Table 1 should satisfy (9).

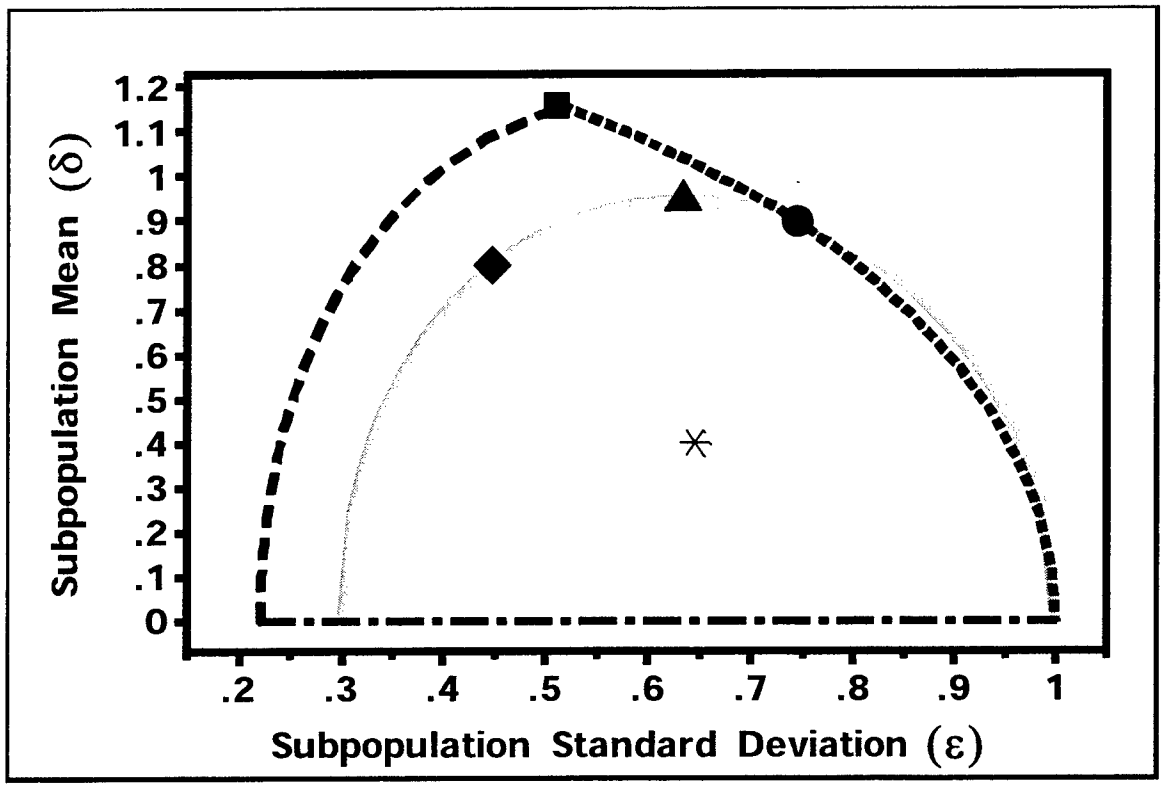

Figure 5. Comparison of Feasible Regions from Covariate and Subpopulation Models

Table 1. Standardized Parameters and the Goodness-of-Fit Criterion for Resistant Subpopulations. For sensitive subpopulations, multiply the means by –1.

	Double Truncation Maximum Mean			Double Truncation Maximum Ratio			Single Truncation		
Size	Mean	StDev	Fit	Mean	StDev	Fit	Mean	StDev	Fit
θ	δ_x	ε_x	C	δ_r	ε_r	C	δ_t	ε_t	C
.05	1.873	.434	.017	.993	.082	.009	1.807	.576	.016
.10	1.554	.489	.017	.974	.163	.012	1.490	.630	.015
.15	1.347	.532	.016	.942	.240	.013	1.287	.667	.014
.20	1.189	.569	.015	.901	.314	.014	1.134	.696	.013
.25	1.059	.602	.015	.853	.383	.014	1.008	.722	.013
.30	.946	.633	.014	.800	.447	.014	.900	.744	.012
.35	.846	.663	.014	.743	.507	.014	.805	.766	.011
.40	.756	.691	.013	.683	.563	.014	.719	.785	.011
.45	.673	.719	.012	.623	.614	.013	.640	.804	.010
.50	.596	.746	.012	.562	.662	.012	.568	.822	.010
.60	.455	.798	.010	.441	.748	.011	.435	.857	.008
.75	.269	.875	.007	.266	.857	.008	.259	.908	.006

In terms of the application, the complementary subgroup consists of all persons who are unfit for military duty. The removal of all healthy, young adults from a population creates an unnatural population in the biological sense; it is not surprising that this unnatural population does not have a normal distribution.

The covariate model only allows for selection by specification of an acceptable range on the covariate. More general selection processes can be modeled as follows. Let $P_Y(y) = 1$ for $a \le Y \le b$ and $P_Y(y) = 0$ otherwise. Then (12) can be written as

$$f_{Xs}(x) = \frac{\int_{-\infty}^{\infty} P_Y(y)\, f_{X,Y}(x,y)\, dy}{\int_{-\infty}^{\infty} P_Y(y)\, f_Y(y)\, dy} \qquad (13)$$

where the integrations extend from $-\infty$ to ∞ mathematically, but, for numerical integration, over an interval sufficient to include nearly all of the distribution. A process for selecting a subpopulation that is more complicated than an acceptable range on a covariate can be incorporated into the covariate model by letting $P_Y(y)$ take on values other than 0 and 1, subject to $0 \le P_Y(y) \le 1$ for all y. Note that $\int P_Y(y)\, f_Y(y)\, dy = \theta$. I call $P_Y(y)$ a selection function. A selection function can also be applied to, or defined for, the variable of interest by using the heights of the histograms in Figure 2: let $P_X(x) = h_s(x) / h_p(x)$.

6. COMPARISON OF METHODS

From (10), the estimate of the ED_{50} for the general population depends on $1/\sigma_s$, the probit slope of the toxicant for the military subpopulation. The probit slopes listed in Grotte and Yang (2001) range from 3 to 12. Therefore, to compare the methods of making estimates, I use fictitious toxicants with an ED_{50} of 100 and probit slopes of 3, 6, and 12 for the military subpopulation. Table 2 gives the ED_{50} and the probit slope for the general population by three methods. The probit slope for the general population is simply ε times the probit slope for the military subpopulation.

Applying a factor of 10 to the parameters used in toxicology, the ED_{50} and the probit slope, yields the estimates $\mu_p = \mu_s - 1$ and $\sigma_p = 10\ \sigma_s$. The uncertainty factor method does not explicitly depend on either the subpopulation size or the subpopulation probit slope.

Table 2. Median Effective Dose and Probit Slope for the General Population

Military Subpopulation			Type of Estimate for the General Population					
Size		Probit Slope	Uncertainty Factor		Double Truncation		Single Truncation	
θ	ED_{50}	$1/\sigma_s$	ED_{50}	$1/\sigma_p$	ED_{50}	$1/\sigma_p$	ED_{50}	$1/\sigma_p$
.2	100	3	10	.3	11	0.9	29	2.1
.2	100	6	10	.6	33	1.9	54	4.2
.2	100	12	10	1.2	57	3.7	73	8.4
.3	100	3	10	.3	26	1.4	39	2.2
.3	100	6	10	.6	51	2.7	63	4.4
.3	100	12	10	1.2	71	5.4	79	8.9
.4	100	3	10	.3	39	1.7	50	2.4
.4	100	6	10	.6	63	3.4	70	4.7
.4	100	12	10	1.2	79	6.7	84	9.5

The estimates for double truncation are based on the maximum-ratio values δ_r and ε_r because the conversion is subpopulation to population. The basis for the use of double truncation is that age is the covariate, and young adults correspond to some age range, say, 18 to 35 years. The straight-line regression of log(effective dose) on age—as implied by the bivariate normal distribution—results in children being more resistant to toxicants than young adults are. Such an assumption is not acceptable for risk assessment. The basis for the use of single truncation is that health status is the covariate, and there is no segment of the population healthier than young adults. Single truncation results in less extreme conversions than double truncation. This result may seem backwards because single truncation puts the subpopulation into the tail of the distribution of the covariate. Single and double truncation are not compared on the same basis because the correlation coefficient ρ is not held constant. To obtain an approximate normal distribution for the variable of interest, single truncation requires a lower value of $|\rho|$ than double truncation does. The lower value of $|\rho|$ results in less difference between the subpopulation and the population. In the application, the ED_{50} and the probit slope for both the population and the subpopulation are used as if they apply to a lognormal distribution, so to compare single and double truncation it seems better to use approximate normal distributions of log(effective dose) for both types of truncation than to use the same value of ρ for both types of truncation.

7. CONCLUSION

As shown in Figure 5, most combinations of parameters indicated as feasible by the subpopulation model are also indicated as feasible by the covariate model. At the combinations of parameters indicated as feasible by the subpopulation

model, the discrepancy between the actual distribution of the subpopulation, as obtained from the covariate model, and the normal distribution assumed by the subpopulation model is small, as indicated by the values of the criterion *C* in Table 1. More extreme worst-case scenarios can be obtained from the covariate model than from the subpopulation model, but only by allowing the subpopulation to have a very non-normal distribution.

The covariate model distinguishes between single and double truncation. The combination of parameters suggested as worst-case values by Crosier (2003) can be obtained from the covariate model only by double truncation on the covariate. However, no covariate for which double truncation is appropriate has been suggested. Age cannot be the covariate because it cannot have a bivariate normal distribution with the logarithm of effective dose—sensitive individuals are at both ends of the age range. If health status is the covariate, then only single truncation is a plausible procedure for selecting military personnel from the general population. Single truncation on the covariate limits the possible parameter combinations to the right-side boundary of the feasible region for the covariate model in Figure 5. Hence, the combination of parameter values previously suggested as the worst-case (the diamond in Figure 5) are not realistic. Similarly, the parameter combinations denoted centroid values (the asterisk in Figure 5) by Crosier (2003) are also not realistic because they require double truncation on the covariate. Another problem with centroid estimates is that they are dependent on the scale over which the averaging is done. For example, they depend on whether the standardization is by the population parameters [$\delta = (\mu_s - \mu_p) / \sigma_p$ and $\varepsilon = \sigma_s / \sigma_p$] or by the subpopulation parameters [$\eta = (\mu_p - \mu_s) / \sigma_s$ and $\psi = \sigma_p / \sigma_s$].

REFERENCES

Cornwell, J.B.; Marx, J.D. The Significance of Hazard End Points in Quantitative Risk Analysis, Quest Consultants, Inc.: Norman, OK (available at http://www.questconsult.com/hazard.html), 2006.

Crosier, R.B.; Sommerville, D.R. *Relationship Between Toxicity Values for the Military Population and Toxicity Values for the General Population*; ECBC-TR-224; U.S. Army Edgewood Chemical Biological Center: Aberdeen Proving Ground, MD, 2002 (AD-A400 214).

Crosier, R.B. *Mathematical Limits on Differences Between a Population and a Subpopulation*; ECBC-TR-337; U.S. Army Edgewood Chemical Biological Center: Aberdeen Proving Ground, MD, 2003 (AD-A417 162).

Dixon, W.J.; Massey, Jr., F.J. Introduction to Statistical Analysis; 3rd ed.; McGraw-Hill: New York, 1969.

Grotte, J. H.; Yang, L.I. *Report of the Workshop on Chemical Agent Toxicity for Acute Effects, Institute for Defense Analyses, May 11-12, 1998*; Institute for Defense Analyses Document D-2176; Institute for Defense Analyses: Alexandria, VA, 2001.

Johnson, Norman L.; Kotz, Samuel Continuous Univariate Distributions-1; John Wiley & Sons: New York, 1970.

Mood, A.M.; Graybill, F.A.; Boes, D.C. Introduction to the Theory of Statistics; 3rd ed.; McGraw-Hill: New York, 1974.

Whalan, J.E.; Foureman, G.L.; Vandenberg, J.J. "Inhalation Risk Assessment at the Environmental Protection Agency," in *Inhalation Toxicology*; 2nd ed.;Salem, H.; Katz, S.A., Eds.; Taylor & Francis: Boca Raton, FL, 2006.

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.